Available online at www.sciencedirect.com
ScienceDirect
Advances in Applied Mathematics 39 (2007) 329–344
ADVANCES IN Applied Mathematics
www.elsevier.com/locate/yaama

On linear matrix differential equations

Luis Verde-Star

Department of Mathematics, Universidad Autónoma Metropolitana, Iztapalapa, Apartado 55-534, México D.F. 09340, Mexico

Received 10 March 2006; accepted 10 June 2006

Available online 15 September 2006

Abstract

We use elementary methods and operator identities to solve linear matrix differential equations and we obtain explicit formulas for the exponential of a matrix. We also give explicit constructions of solutions of scalar homogeneous equations with certain initial values, called dynamic solutions, that play an important role in the solution of homogeneous and non-homogeneous matrix differential equations. We show that the same methods can be used to solve linear matrix difference equations.

MSC: 34A30; 34A05; 39A10

Keywords: Matrix differential equations; Matrix exponential; Exponential polynomials; Dynamic solution; Convolutions; Matrix difference equations

1. Introduction

We consider matrix differential equations of the form

$$M'(t) = AM(t) + U(t), \quad t \in \mathbb{C}, \tag{1.1}$$

where A is a constant square matrix, $U(t)$ is a given matrix function, and $M(t)$ is an unknown matrix function. These equations appear often in many areas of mathematics and its applications, and finding their solutions is an important computational problem. The homogeneous case of (1.1) leads to the matrix exponential e^{tA}. This function is used to construct particular solutions of the non-homogeneous equations by means of the convolution product of functions.

E-mail address: verde@xanum.uam.mx.

0196-8858/$ – see front matter
doi:10.1016/j.aam.2006.06.002

We present some simple methods for the construction of solutions of (1.1). We use the fact that finding a solution of the homogeneous matrix equation (1.1) reduces to finding a single solution $g(t)$ of a scalar differential equation of the form $w(D)g = 0$, where w is any monic polynomial such that $w(A) = 0$. The solution e^{tA} corresponds to a solution g with certain initial values, called the dynamic solution. We present explicit constructions for the dynamic solution. Higher order matrix differential equations are also considered.

We also show that the same methods can be used to solve the discrete version of (1.1)

$$M(k+1) = AM(k) + U(k), \quad k \geqslant 0, \tag{1.2}$$

which is a matrix difference equation, and leads to the sequence of powers A^k in the homogeneous case.

2. Construction of the matrix exponential

Let m be a positive integer. We denote by $\mathcal{M}$ the complex vector space of $m \times m$ matrix-valued functions of the complex variable t. Let $M(t)$ be an element of $\mathcal{M}$ and let A be a constant $m \times m$ matrix. We consider first the homogeneous equation

$$M'(t) = AM(t). \tag{2.1}$$

It is well-known that the solution of this equation that satisfies $M(0) = I$ is the matrix exponential function e^{tA}, usually defined by the series

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k. \tag{2.2}$$

Let

$$w(t) = b_0 t^{n+1} + b_1 t^n + b_2 t^{n-1} + \cdots + b_{n+1} \tag{2.3}$$

be any polynomial, with $b_0 = 1$, such that $w(A) = 0$. It is clear that w may be any monic polynomial that is divisible by the minimal polynomial of A. By the Cayley–Hamilton theorem, $w(t) = \det(tI - A)$, the characteristic polynomial of A, is an acceptable w.

The equation $w(A) = 0$ yields

$$A^{n+1} = -\sum_{j=0}^{n} b_{j+1} A^{n-j}, \tag{2.4}$$

and thus

$$A^{n+1+k} = -\sum_{j=0}^{n} b_{j+1} A^{n+k-j}, \quad k \geqslant 0. \tag{2.5}$$

Therefore all the powers A^k, for $k \geqslant n+1$, can be written as a linear combination of $I, A, A^2, \dots, A^n$. This suggests that the series for $M(t)$ could be written in the form

$$M(t) = \sum_{k=0}^{n} f_k(t) A^{n-k}. \tag{2.6}$$

Suppose that (2.1) has a solution of this form. In order to find the coefficient functions $f_k(t)$ we use first a method based on Putzer's ideas [6], and another simple procedure based on operator identities.

Substitution of (2.6) in (2.1) gives

$$\sum_{k=0}^{n} f_k'(t) A^{n-k} = f_0(t) A^{n+1} + \sum_{k=1}^{n} f_k(t) A^{n+1-k}. \tag{2.7}$$

Combining this equation with (2.4) and collecting terms, we get

$$\big\{ f_n'(t) + b_{n+1} f_0(t) \big\} I + \sum_{k=0}^{n-1} \big\{ f_k'(t) - f_{k+1}(t) + b_{k+1} f_0(t) \big\} A^{n-k} = 0. \tag{2.8}$$

For this equation to hold it is sufficient that the coefficients of the powers of A be zero. Note that the matrices $I, A, A^2, \dots, A^n$ need not be linearly independent. Then we look for functions f_k that satisfy the $n+1$ equations

$$f_{k+1}(t) - f_k'(t) = b_{k+1} f_0(t), \quad 0 \leqslant k \leqslant n-1, \tag{2.9}$$

and

$$-f_n'(t) = b_{n+1} f_0(t). \tag{2.10}$$

Differentiating the kth equation $n-k$ times in (2.9) we obtain

$$f_{k+1}^{(n-k)}(t) - f_k^{(n+1-k)}(t) = b_{k+1} f_0^{(n-k)}(t), \quad 0 \leqslant k \leqslant n-1. \tag{2.11}$$

Adding all these equations with (2.10) we obtain a telescopic sum that yields

$$\sum_{k=0}^{n+1} b_k D^{n+1-k} f_0(t) = 0. \tag{2.12}$$

This equation can be written in the form $w(D) f_0(t) = 0$, which is a scalar homogeneous linear differential equation with constant coefficients.

Equation (2.9) can be written as

$$f_{k+1}(t) = f_k'(t) + b_{k+1} f_0(t), \quad 0 \leqslant k \leqslant n-1, \tag{2.13}$$

and by repeated substitution yields

$$f_k(t) = \left(D^k + b_1 D^{k-1} + b_2 D^{k-2} + \cdots + b_k\right) f_0(t), \quad 1 \leqslant k \leqslant n, \tag{2.14}$$

that gives us $f_1, f_2, \ldots, f_n$ in terms of f_0. Therefore, finding a solution of (2.1) of the form (2.6) is reduced to finding a solution of the scalar equation $w(D)g(t) = 0$.

From (2.6) it is clear that we can have $M(0) = I$ if $f_n(0) = 1$ and $f_k(0) = 0$ for $0 \leqslant k \leqslant n-1$. In view of (2.14), these initial values for the f_k are obtained if f_0 satisfies $D^k f_0(0) = 0$ for $0 \leqslant k \leqslant n-1$ and $D^n f_0(0) = 1$. The solution f_0 of $w(D)g(t) = 0$ with such initial values is called the *dynamic solution* associated with w.

We summarize the previous results as follows. Given a square matrix A and a monic polynomial w such that $w(A) = 0$, the solution $M(t)$ of (2.1) that satisfies $M(0) = I$ is given by (2.6), where the coefficients f_k are given by (2.14), the b_k are the coefficients of w, and f_0 is the dynamic solution associated with w.

Notice that (2.14) shows that $f_1, f_2, \ldots, f_n$ are also solutions of $w(D)g = 0$ and, in view of (2.6), all the entries of $M(t)$ are also solutions of such equation. Therefore $M(t)$ satisfies $w(D)M(t) = 0$.

Before we give a direct construction of the dynamic solution f_0 in the next section, we present another approach to the construction of the matrix exponential. Note that (2.13) is a sort of differential Horner's algorithm and that (2.14) describes the f_k as the intermediate results in a sort of "synthetic division" computation that ends with $f_{n+1}(t) = w(D)f_0(t) = 0$. This is our motivation for the following development.

Let $w(t)$ be a monic polynomial of degree $n+1$ with complex coefficients as in (2.3). Define the difference quotient $w[t,x]$ by

$$w[t,x] = \frac{w(t) - w(x)}{t - x}.$$

The basic polynomial identity

$$t^{k+1} - x^{k+1} = (t-x)\sum_{j=0}^{k} t^j x^{k-j}, \quad k \geqslant 0, \tag{2.15}$$

clearly shows that $w[t,x]$ is a symmetric polynomial in t and x of degree n in each variable. Thus

$$(t-x)w[t,x] = w(t) - w(x) \tag{2.16}$$

is a polynomial identity.

Let D denote differentiation with respect to the complex variable t, and let A be an $m \times m$ matrix. Note that, on the elements of $\mathcal{M}$, the operator D acts entrywise, A acts by multiplication on the left, and the operators D and A commute.

Replacing t by D and x by A in $w[t,x]$ we obtain a symmetric polynomial $w[D,A]$, which is a linear differential operator of order n with constant matrix coefficients. Since (2.16) is a polynomial identity and D and A commute, $w[D,A]$ satisfies the operator identity

$$(D-A)w[D,A] = w(D) - w(A). \tag{2.17}$$

Suppose now that A satisfies $w(A)=0$ and let $f(t)$ be any scalar function such that $w(D)f(t)=0$. Apply the operators that appear in (2.17) to $f(t)I$, where I denotes the identity matrix, to obtain

$$(D-A)w[A,D]f(t)I = w(D)f(t)I - w(A)f(t)I = 0. \tag{2.18}$$

Therefore the matrix function $M(t)=w[A,D]f(t)I$ is a solution of (2.1). Note that $M(t)$ is a polynomial in A of degree n with coefficients that depend on t. It is also a linear combination of the functions $D^k f(t)$, for $0 \leqslant k \leqslant n$, with coefficients that are polynomials in A with degree less than or equal to n.

Using (2.15) we obtain

$$w[t,x] = \sum_{k=0}^{n} b_k \sum_{j=0}^{n-k} t^j x^{n-k-j} = \sum_{j=0}^{n} \left\{ \sum_{k=0}^{n-j} b_k x^{n-k-j} \right\} t^j = \sum_{j=0}^{n} w_{n-j}(x) t^j, \tag{2.19}$$

where the polynomials w_k are defined by

$$w_k(x) = b_0 x^k + b_1 x^{k-1} + \cdots + b_k, \quad 0 \leqslant k \leqslant n+1. \tag{2.20}$$

Note that

$$w_{k+1}(x) = x w_k(x) + b_{k+1}, \quad 0 \leqslant k \leqslant n, \tag{2.21}$$

and $w_{n+1}(x)=w(x)$. We call the w_k the *Horner polynomials* of w because (2.21) is Horner's algorithm for the computation of $w(x)$. It is clear that $\{w_k\colon\ 0 \leqslant k \leqslant n\}$ is a basis for the complex vector space of polynomials of degree less than or equal to n. This basis is also called the *control basis* associated with w.

Using (2.19) and the symmetry of $w[t,x]$ we can write $w[A,D]$ as follows

$$w[A,D] = \sum_{j=0}^{n} A^{n-j} w_j(D) = \sum_{j=0}^{n} w_{n-j}(A) D^j. \tag{2.22}$$

Therefore, for $M(t)=w[A,D]f(t)I$ we have the expressions

$$M(t) = \sum_{j=0}^{n} A^{n-j} w_j(D) f(t) = \sum_{j=0}^{n} w_{n-j}(A) D^j f(t). \tag{2.23}$$

The last expression shows that taking suitable values for $D^k f(0)$, for $0 \leqslant k \leqslant n$, we can obtain any polynomial in A as the value of $M(0)$. In particular, if $f(t)$ is the dynamic solution associated with w then $M(0)=I$, and hence $M(t)=e^{tA}$. We summarize the above results in the following.

Theorem 2.1. *Let w be a monic polynomial and let $w_0, w_1, \dots, w_n$ be the Horner polynomials of w. Let $f(t)$ be the dynamic solution associated with w. Let A be any square matrix such that $w(A) = 0$. Then we have*

$$e^{tA} = \sum_{j=0}^{n} A^{n-j} w_j(D) f(t), \tag{2.24}$$

and

$$e^{tA} = \sum_{j=0}^{n} w_j(A) D^{n-j} f(t). \tag{2.25}$$

Note that the dependence on t of e^{tA} is completely determined by the dynamic solution $f(t)$, which is in turn determined by the polynomial w. Note also that we can use the recurrence relation (2.21) to compute the summands in (2.24) and (2.25).

3. The dynamic solution associated with w

We present first a simple construction for the coefficients of the power series representation of the dynamic solution associated with w. An explicit expression for the dynamic solution as an exponential polynomial will be given later.

Let w be given by (2.3). Define the reversed polynomial $w^*(t) = t^{n+1} w(1/t)$. Note that

$$w^*(t) = 1 + b_1 t + b_2 t^2 + \cdots + b_{n+1} t^{n+1}. \tag{3.1}$$

Let the formal power series $h(t) = \sum_{k \geqslant 0} h_k t^k$ be the reciprocal of $w^*(t)$. That is, $h(t) w^*(t) = 1$. Then the coefficients h_k satisfy $h_0 = 1$ and

$$h_k b_0 + h_{k-1} b_1 + \cdots + h_0 b_k = 0, \quad k \geqslant 1. \tag{3.2}$$

Since $b_0 = 1$, this is a linear recurrence relation that can be used to compute all the h_k recursively. Using Cramer's rule on the system of linear equations that consists of $h_0 = 1$ and the equations obtained from (3.2) by taking $1 \leqslant k \leqslant j$, we obtain

$$h_j = (-1)^j \det \begin{bmatrix} b_1 & 1 & & & \\ b_2 & b_1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ b_{j-1} & b_{j-2} & b_{j-3} & \dots & 1 \\ b_j & b_{j-1} & b_{j-2} & \dots & b_1 \end{bmatrix}, \quad j \geqslant 1. \tag{3.3}$$

Recall that $b_j = 0$ if $j > n + 1$ and thus the matrix above is banded if $j > n + 1$. This determinant representation for the h_j shows the complexity of the dependence of h_j on the coefficients $b_1, b_2, \dots, b_j$, but it is not a practical way to compute the h_j.

Define

$$f_w(t) = \sum_{k=0}^{\infty} h_k \frac{t^{k+n}}{(k+n)!}. \tag{3.4}$$

We claim that $f_w(t)$ is the dynamic solution associated with w. The convergence of (3.4) for all complex values of t is a consequence of another expression for $f_w(t)$ that we will give later. An alternative proof is obtained by using the fact that the h_k are the coefficients of the Taylor series of the reciprocal of a polynomial, and that such series converge outside some disk.

Let us show first that $w(D)f_w = 0$. Since $D^j(t^k/k!) = t^{k-j}/(k-j)!$ for $j \leqslant k$, and $D^j(t^k/k!) = 0$ for $j > k$, by a simple rearrangement of terms it is easy to see that

$$\begin{aligned} w(D)f_w(t) &= \sum_{k\geqslant 0}\sum_{j=0}^{n+1} h_k b_{n+1-j}\frac{t^{n+k-j}}{(n+k-j)!} \\ &= \sum_{k\geqslant 0}\sum_{\ell=k-1}^{n+k} h_k b_{\ell+1-k}\frac{t^\ell}{\ell!} \\ &= \sum_{\ell\geqslant 0}\sum_{k=\ell-n}^{\ell+1} h_k b_{\ell+1-k}\frac{t^\ell}{\ell!} = 0. \end{aligned}$$

The last equality follows from (3.2), which shows that all the coefficients in the last power series are zero. Note that in the above sums we take $h_k = 0$ whenever $k < 0$.

Since t^n is a common factor of all the terms in the series for $f_w(t)$, and the coefficient of the first term is $h_0 = 1$, it is clear that $f_w(t)$ has the correct initial values to be the dynamic solution associated with w.

Any Taylor polynomial approximation of $f_w(t)$ obtained by truncation of (3.4) yields immediately a polynomial approximation for e^{tA}, by substitution of f_w by its approximating polynomial in (2.24) or (2.25).

We find next another representation for the dynamic solution. Let

$$w(t) = \prod_{j=0}^{r}(t-a_j)^{m_j+1}, \tag{3.5}$$

where the a_j are distinct complex numbers and the m_j are nonnegative integers such that $\sum_j(m_j+1) = n+1$. We consider first the differential operators of the form $(D-aI)^k$, because $w(D)$ is a product of a finite number of such operators. The *basic exponential polynomials* are the functions

$$g_{a,k}(t) = \frac{t^k}{k!}e^{at}, \quad a \in \mathbb{C},\ k \in \mathbb{N}, \tag{3.6}$$

defined on the complex plane. It is easy to show that these functions are linearly independent. Denote by $\mathcal{E}$ the complex vector space generated by all the $g_{a,k}$. Its elements are called exponential polynomials.

A simple computation gives

$$(D-aI)^j g_{a,k}(t) = \begin{cases} g_{a,k-j}(t), & k \geqslant j \\ 0, & k < j. \end{cases} \tag{3.7}$$

In particular, any linear combination of the $g_{a,k}(t)$, for $0 \leqslant k < j$, is an element of the null space of $(D - aI)^j$. Therefore, the vector space generated by the set $S = \{g_{a_j,k}\colon 0 \leqslant j \leqslant r,\ 0 \leqslant k \leqslant m_j\}$ is contained in the null space of $w(D)$. In fact, it is easy to see that S is a basis for the null space of $w(D)$.

We define a commutative multiplication $*$ on $\mathcal{E}$, called the convolution product, as follows. For $a \neq b$

$$g_{a,m} * g_{b,k} = \sum_{j=0}^{m} C(a, j, b, k) g_{a,m-j} + \sum_{j=0}^{k} C(b, j, a, m) g_{b,k-j}, \tag{3.8}$$

where the coefficients are defined by

$$C(a, j, b, i) = (-1)^j \binom{j+i}{j} (a-b)^{-j-i-1}, \quad a \neq b,\ i, j \in \mathbb{N}, \tag{3.9}$$

and for the remaining case define

$$g_{a,k} * g_{a,m} = g_{a,1+k+m}, \quad a \in \mathbb{C},\ k, m \in \mathbb{N}. \tag{3.10}$$

Let Φ be the linear functional on $\mathcal{E}$ defined by $\Phi g_{a,k} = \delta_{0,k}$, for (a, k) in $\mathbb{C} \times \mathbb{N}$. Note that Φ can be seen as evaluation at $t = 0$. A straightforward computation gives

$$D(g_{a,k} * g_{b,m}) = (D g_{a,k}) * g_{b,m} + g_{b,m} \Phi g_{a,k}, \tag{3.11}$$

and by linearity we have

$$D(g * f) = (Dg) * f + f \Phi g, \quad f, g \in \mathcal{E}. \tag{3.12}$$

Theorem 3.1. *Let $w(t)$ be as in* (3.5) *and let*

$$f_w = g_{a_0,m_0} * g_{a_1,m_1} * \cdots * g_{a_r,m_r}. \tag{3.13}$$

Then f_w is the dynamic solution associated with w.

Proof. From the definition of the convolution it is clear that the subspace of $\mathcal{E}$ generated by S is closed under the convolution product and thus f_w is in the null space of $w(D)$.

From (3.8) and (3.9) we get

$$\Phi(g_{a,m} * g_{b,k}) = C(a, m, b, k) + C(b, k, a, m) = 0$$

for $a \neq b$ and any m and k in $\mathbb{N}$, and from (3.10)

$$\Phi(g_{a,k} * g_{a,m}) = \Phi g_{a,k+m+1} = 0, \quad k, m \in \mathbb{N}.$$

Therefore the value at zero of the convolution product of any two basic exponential polynomials is zero. Thus we have

$$\Phi(g_{a,k} * f) = 0, \quad (a, k) \in \mathbb{C} \times \mathbb{N},\ f \in \mathcal{E}. \tag{3.14}$$

We determine next the initial values of f_w. Consider first the case when w has only one root a of multiplicity $n+1$ and let $0 \leqslant k \leqslant n$. Then $f_w = g_{a,n}$ and writing $D^k = ((D-aI)+aI)^k$ we get

$$D^k f_w = \sum_{\ell=0}^{k} \binom{k}{\ell} a^{k-\ell}(D-aI)^\ell g_{a,n} = \sum_{\ell=0}^{k} \binom{k}{\ell} a^{k-\ell} g_{a,n-\ell}$$

and thus

$$\Phi D^k f_w = \sum_{\ell=0}^{k} \binom{k}{\ell} a^{k-\ell} \delta_{\ell,n} = \delta_{k,n}.$$

Suppose now that w has at least two distinct roots, let $n > 1$ and let $0 \leqslant k \leqslant n$. Using (3.10) we can write $f_w = g * h$ where $g = g_{a_0,0}$ and h is the convolution of the g_{a_i,m_i} for $i > 0$ and also g_{a_0,m_0-1} if $m_0 > 0$. Note that $h = f_u$ where $u(t) = w(t)/(t-a_0)$ and $u(t)$ has degree equal to n. This fact is used in the induction step below. Note also that $\Phi D^j g = a_0^j$ for $j \geqslant 0$. Then $D(g*h) = Dg*h + h\Phi g$ and by iteration we obtain

$$D^k(g*h) = D^k g * h + \sum_{\ell=0}^{k-1} a_0^{k-1-\ell} D^\ell h.$$

Applying Φ to both sides of this equation and using (3.14) and induction on n we get

$$\Phi D^k(g*h) = \sum_{\ell=0}^{k-1} a_0^{k-1-\ell} \delta_{\ell,n-1} = \delta_{k-1,n-1} = \delta_{k,n}.$$

Therefore $f_w = g*h$ has the correct initial values to be the dynamic solution associated with w. □

By the uniqueness theorem for the solutions of linear differential equations with constant coefficients we see that the exponential polynomial f_w defined above must coincide with the power series f_w defined in (3.4).

The computation of f_w using (3.13) is a straightforward repeated application of the convolution formula (3.8). It is in fact essentially equivalent to finding the partial fractions decomposition of $1/w(t)$. See [12].

A simple computation gives us

$$(D-aI)^{k+1}(g_{a,k} * h) = h, \quad (a,k) \in \mathbb{C} \times \mathbb{N},\ h \in \mathcal{E}. \tag{3.15}$$

This means that the map that sends h to $g_{a,k} * h$ is a right inverse for $(D-aI)^{k+1}$. Applying this result repeatedly we obtain the following.

Theorem 3.2. *Let $w(t)$ be a polynomial as in* (3.5) *and let f_w be given by* (3.13). *Then the linear map on $\mathcal{E}$ that sends h to $f_w * h$ is a right inverse for the operator $w(D)$.*

The ideas used in the proof of Theorem 3.1 can be used to show that the particular solution $f_w * h$ of $w(D)g = h$ has all its initial values equal to zero. See [12].

4. Non-homogeneous matrix equations

Let us consider now the non-homogeneous matrix equation

$$M'(t) = AM(t) + U(t), \tag{4.1}$$

where A is a constant $m \times m$ matrix, $U(t)$ is a given element of $\mathcal{M}$ and $M(t)$ is an unknown matrix function. Let

$$U(t) = \sum_{j=0}^{s} \alpha_j(t) B_{s-j}, \tag{4.2}$$

where the B_j are constant $m \times m$ matrices and suppose that the $\alpha_j(t)$ are exponential polynomials.

Let $w(t) = t^{n+1} + b_1 t^n + \cdots + b_{n+1}$ be a monic polynomial such that $w(A) = 0$, let $f_0(t)$ be the dynamic solution associated with w, and let $f_k(t)$ be defined by (2.14). Then, by the results of Section 2 we have

$$e^{tA} = \sum_{k=0}^{n} f_k(t) A^{n-k}.$$

We define the convolution

$$e^{tA} * U(t) = \sum_{k=0}^{n} \sum_{j=0}^{s} \big(f_k(t) * \alpha_j(t)\big) A^{n-k} B_{s-j}. \tag{4.3}$$

Note that this convolution is not commutative because of the noncommutativity of the matrix multiplication.

Theorem 4.1. *The matrix function $e^{tA} * U(t)$ is a particular solution of* (4.1).

Proof. By (3.12) we have

$$D\big(f_k(t) * \alpha_j(t)\big) = f_k'(t) * \alpha_j(t) + \alpha_j(t) f_k(0) \tag{4.4}$$

and $f_k(0) = \delta_{k,n}$ by (2.14) because f_0 is the dynamic solution. Therefore

$$D\big(e^{tA} * U(t)\big) = \sum_{k=0}^{n} \sum_{j=0}^{s} \big(f_k'(t) * \alpha_j(t)\big) A^{n-k} B_{s-j} + \sum_{k=0}^{n} \sum_{j=0}^{s} \delta_{k,n} \alpha_j(t) A^{n-k} B_{s-j}.$$

The last sum is clearly equal to $U(t)$. Since e^{tA} is a solution of $M'(t) = AM(t)$ the first sum equals $Ae^{tA} * U(t)$. This completes the proof. $\square$

An apparent limitation of the previous theorem is the hypothesis that the functions $\alpha_j(t)$ must be exponential polynomials. We can eliminate this limitation as follows. Note that the proof of the theorem only uses the fact that the convolutions $f_k(t) * \alpha_j(t)$ are well defined and that (4.4) holds.

The *Duhamel convolution* of functions f and g of a real or complex variable t is defined by

$$f * g(t) = \int_0^t f(t-z)g(z)\,dz. \tag{4.5}$$

Using elementary tools it is easy, but quite laborious, to show that the convolution defined in (3.8) and the Duhamel convolution coincide on the space $\mathcal{E}$ of exponential polynomials. We can allow the $\alpha_j(t)$ to be any functions for which their Duhamel convolution with any basic exponential polynomial is well defined. For example, they can be piecewise continuous real functions. Using basic properties of integration it is easy to show that (4.4) also holds in this case.

5. Higher order equations

We consider now linear matrix differential equations of order r with matrix coefficients. Let $B_0, B_1, \ldots, B_r$ be $m \times m$ matrices, with $B_0 = I$. Define the differential operator

$$L = \sum_{j=0}^{r} B_j D^{r-j}, \tag{5.1}$$

where D denotes differentiation with respect to t. We want to find a matrix function $F(t)$ such that

$$LF(t) = \sum_{j=0}^{r} B_j F^{(r-j)}(t) = 0, \tag{5.2}$$

for $t \in \mathbb{C}$, and

$$F^{(j)}(0) = Q_j, \quad 0 \leqslant j \leqslant r-1, \tag{5.3}$$

where the Q_j are given matrices.

Define the block companion matrix

$$A = \begin{bmatrix} 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & I \\ -B_r & -B_{r-1} & -B_{r-2} & \cdots & -B_2 & -B_1 \end{bmatrix}, \tag{5.4}$$

where the entries are $m \times m$ matrices. Let $M(t) = e^{tA}$. Then we have $M'(t) = AM(t)$ and $M(0) = I$. We consider $M(t)$ as an $r \times r$ matrix whose entries $M_{i,j}(t)$ are $m \times m$ matrices. Since $M'(t) = AM(t)$, it is easy to see that

$$M_{i,k}(t) = M_{1,k}^{(i-1)}(t), \quad 1 \leqslant i \leqslant r, \ 1 \leqslant k \leqslant r. \tag{5.5}$$

and

$$LM_{1,k}(t) = 0, \quad 1 \leqslant k \leqslant r. \tag{5.6}$$

Since $M(t)$ commutes with A we have $M'(t) = M(t)A$, and looking at the first row in this matrix equation we get

$$M'_{1,k}(t) = M_{1,k-1}(t) - M_{1,r}(t)B_{r+1-k}, \quad 1 \leqslant k \leqslant r, \tag{5.7}$$

where $M_{1,0}(t) = 0$. Solving for $M_{1,k-1}(t)$ and taking $k = r, r-1, \ldots, r-j+1$ we obtain

$$\begin{aligned}
M_{1,r-1}(t) &= M'_{1,r}(t) + M_{1,r}(t)B_1, \\
M_{1,r-2}(t) &= M'_{1,r-1}(t) + M_{1,r}(t)B_2, \\
\vdots \qquad & \qquad \vdots \qquad\qquad \vdots \\
M_{1,r-j}(t) &= M'_{1,r-j+1}(t) + M_{1,r}(t)B_j.
\end{aligned}$$

By repeated differentiation and forward substitution we obtain

$$M_{1,r-j}(t) = M_{1,r}^{(j)}(t) + M_{1,r}^{(j-1)}(t)B_1 + \cdots + M_{1,r}(t)B_j, \quad 1 \leqslant j \leqslant r-1. \tag{5.8}$$

We denote by $L_j M_{1,r}(t)$ the expression in the right-hand side.

Let

$$F(t) = \sum_{k=1}^{r} Q_{k-1} M_{1,k}(t) = \sum_{k=1}^{r} Q_{r-k} L_{k-1} M_{1,r}(t). \tag{5.9}$$

From (5.6) we see that $F(t)$ satisfies (5.2). Since $M(0) = I$, by (5.5) we have

$$M_{i,k}(0) = \delta_{i,k} I = M_{1,k}^{(i-1)}(0), \quad 1 \leqslant i \leqslant r, \ 1 \leqslant k \leqslant r.$$

Therefore $F(t)$ satisfies (5.3). Note that $F(t)$ is expressed in terms of the matrix function $M_{1,r}(t)$. This is a solution of (5.2) with initial values

$$M_{1,r}^{(r-1)}(0) = I, \qquad M_{1,r}^{(j)}(0) = 0, \quad 0 \leqslant j \leqslant r-2.$$

We call $M_{1,r}(t)$ the *dynamic solution* of (5.2).

Since $M_{1,r}(t)$ is a block of e^{tA}, we can use the results of the previous sections to find explicit formulas for its computation.

Let w be the characteristic polynomial of A. Let $n = rm - 1$ and f be the dynamic solution associated with w. Then, by (2.24) we have

$$M(t) = \sum_{j=0}^{n} w_{n-j}(D) f(t) A^j. \tag{5.10}$$

Since A is a block companion matrix, if we write the last column of A^j in the form $[P_j\ P_{j+1}\ \cdots\ P_{j+r-1}]^{\mathsf{T}}$, for $j \geqslant 0$, then the sequence P_j of $m \times m$ matrices satisfies the linear recurrence relation

$$P_{j+r} = -\sum_{\ell=1}^{r} B_\ell P_{j+r-\ell}, \quad j \geqslant 0, \tag{5.11}$$

and has initial values $P_j = 0$ for $0 \leqslant j \leqslant r-2$, and $P_{r-1} = I$. The sequence P_j is called the dynamic solution of (5.11).

Therefore, the block in the $(1, r)$ position of $M(t)$ is

$$M_{1,r}(t) = \sum_{j=0}^{n} w_{n-j}(D) f(t) P_j. \tag{5.12}$$

Note that $M_{1,r}(t)$ is completely determined by the characteristic polynomial of A and the first rm terms of the sequence P_j, which appear in the last column of the block decomposition of $I, A, A^2, \ldots, A^n$.

Combining (5.12) and (5.9) we obtain

$$F(t) = \sum_{k=1}^{r} \sum_{j=0}^{n} Q_{r-k} L_{k-1}\big(w_{n-j}(D) f(t) P_j\big). \tag{5.13}$$

A similar result was obtained in [7], but using methods based on the Laplace–Stieltjes transform.

6. Matrix difference equations

We present here some examples of matrix difference equations that can be solved with the methods presented in the previous sections.

For any scalar or matrix-valued function $f(\ell)$, defined for ℓ in the set $\mathbb{N}$ of nonnegative integers, the *forward shift operator* E is defined by $Ef(\ell) = f(\ell + 1)$. Let A be a given $m \times m$ matrix. Consider the matrix difference equation

$$M(\ell + 1) = AM(\ell), \quad \ell \geqslant 0. \tag{6.1}$$

It is obvious that the solution of this equation is $M(\ell) = A^\ell M(0)$.

From the polynomial identity (2.16) we get the operator identity

$$(E - A)w[E, A] = w(E) - w(A). \tag{6.2}$$

Suppose now that $w(A)=0$ and let $g(\ell)$ be any scalar solution of the difference equation $w(E)f(\ell)=0$. Then, the operator identity (6.2) shows that the matrix function $M(\ell)=w[E,A]g(\ell)I$ is a solution of (6.1). It is easy to see that if g is the solution with initial values $g(\ell)=0$ for $0\leqslant\ell\leqslant n-1$ and $g(n)=1$, then $M(0)=I$ and $M(\ell)=A^\ell$ for all $\ell\geqslant 0$. Such sequence g is called the dynamic solution of $w(E)f(\ell)=0$.

Using the ideas introduced in Section 2 we obtain immediately the following results.

Theorem 6.1. *Let w be a monic polynomial and let $w_0, w_1, \dots, w_n$ be the Horner polynomials of w. Let $g(\ell)$ be the dynamic solution of the scalar difference equation $w(E)g=0$. Let A be any square matrix such that $w(A)=0$. Then we have*

$$A^\ell=\sum_{k=0}^{n} A^{n-k} w_k(E) g(\ell), \quad \ell\geqslant 0, \tag{6.3}$$

and

$$A^\ell=\sum_{k=0}^{n} g(\ell+n-k) w_k(A), \quad \ell\geqslant 0. \tag{6.4}$$

We find next a simple explicit construction of the dynamic solution $g(\ell)$. Define the *basic linearly recurrent sequences*

$$s_{a,k}(\ell)=\binom{\ell}{k} a^{\ell-k}, \quad a\in\mathbb{C},\ k\in\mathbb{N}. \tag{6.5}$$

Let $\mathcal{S}$ denote the complex vector space generated by the sequences $s_{a,k}(\ell)$. The elements of $\mathcal{S}$ are called linearly recurrent sequences.

Note that

$$(E-aI)^j s_{a,k}(\ell)=\begin{cases} s_{a,k-j}(\ell), & k\geqslant j\\ 0, & k<j. \end{cases} \tag{6.6}$$

In particular, any linear combination of the sequences $s_{a,k}(\ell)$ for $0\leqslant k<j$ is an element of the null space of $(E-aI)^j$. Therefore, if $w(t)$ is as in (3.5) then the vector space generated by $\{s_{a_j,k}\colon\ 0\leqslant j\leqslant r,\ 0\leqslant k\leqslant m_j\}$ is contained in the null space of $w(E)$.

We define the convolution product $*$ on $\mathcal{S}$ as follows. For $a\neq b$

$$s_{a,m} * s_{b,k}=\sum_{j=0}^{m} C(a,j,b,k) s_{a,m-j}+\sum_{j=0}^{k} C(b,j,a,m) s_{b,k-j}, \tag{6.7}$$

where the coefficients $C(a,j,b,i)$ are defined in (3.9), and $s_{a,m} * s_{a,k}=s_{a,1+m+k}$.

Notice that the previous statements are obtained by putting E in the place of D, $s_{a,k}$ in the place of $g_{a,k}$, and ℓ in the place of t in some statements from Section 3. We are using here a linear isomorphism between the spaces $\mathcal{E}$ and $\mathcal{S}$ determined by the correspondence $g_{a,k}\to s_{a,k}$. The operators E and D are similar in the linear algebraic sense. In [12] a linear algebra approach is used to study functional equations determined by an operator L that is similar to the differentiation operator D.

The results of Section 3 about D and the convolution on $\mathcal{E}$ are transformed immediately to results about E and the convolution on $\mathcal{S}$. For example, from Theorem 3.1 we obtain the following.

Theorem 6.2. *Let $w(t)$ be as in* (3.5) *and let*

$$f_w = s_{a_0,m_0} * s_{a_1,m_1} * \cdots * s_{a_r,m_r}. \tag{6.8}$$

Then $f_w(\ell)$ is the dynamic solution associated with w.

In the same way, we can imitate the results of Section 4 to find particular solutions of the non-homogeneous difference equation

$$M(\ell+1) = AM(\ell) + U(\ell), \tag{6.9}$$

where $U(\ell)$ is a given matrix sequence. If

$$U(\ell) = \sum_{j=0}^{s} \alpha_j(\ell) B_{s-j}, \tag{6.10}$$

where the B_j are constant $m \times m$ matrices and the $\alpha_j(\ell)$ are arbitrary sequences, and A^ℓ is given by (6.4) then

$$A^\ell * U(\ell) = \sum_{k=0}^{n} \sum_{j=0}^{s} \big(g(\ell+n-k) * \alpha_j(\ell)\big) w_k(A) B_{s-j} \tag{6.11}$$

is a particular solution of (6.9).

The extension to the case of higher order difference equations can be done using the method presented in the previous section for differential equations. See [11].

7. Final remarks

In [12] we motivate the introduction of the convolution product and arrive at its definition using explicit expressions for right inverses of simple differential operators. We also explain the connection of the convolution with multiplication of proper rational functions and partial fractions decomposition, and show that our methods can be used to solve other types of linear functional equations, including some differential and difference equations with variable coefficients.

The idea of using the dynamic solution to construct solutions of matrix equations was used in [7,8]. For the numerical aspects of the computation of the matrix exponential see [5]. Some interesting ideas about the exponential and the powers of a matrix are presented in [1–4,6,9].

In [10,11] other types of matrix difference and differential equations are solved by methods similar to the ones used in the present paper. Another approach to the study of functions of matrices is presented in [13].

References

[1] S.N. Elaydi, W.A. Harris Jr., On the computation of A^n, SIAM Rev. 40 (1998) 965–971.
[2] R.B. Kirchner, An explicit formula for e^{tA}, Amer. Math. Monthly 74 (1967) 1200–1204.
[3] I.E. Leonard, The matrix exponential, SIAM Rev. 38 (1996) 507–512.
[4] E. Liz, A note on the matrix exponential, SIAM Rev. 40 (1998) 700–702.
[5] C.B. Moler, C.F. Van Loan, Nineteen dubious ways of computing the exponential of a matrix, twenty-five years later, SIAM Rev. 45 (2003) 3–49.
[6] E.J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, Amer. Math. Monthly 73 (1966) 2–7.
[7] J.C. Ruiz Claeyssen, E. Gallicchio, M.T. Vilhena, Inversion of higher-order matrix difference and differential equations through their dynamical solutions, J. Math. Anal. Appl. 149 (1990) 369–376.
[8] J.C. Ruiz Claeyssen, T. Tsukazan, Dynamic solutions of linear matrix differential equations, Quart. Appl. Math. 48 (1990) 169–179.
[9] H.-J. Runckel, U. Pittelkow, Practical computation of matrix functions, Linear Algebra Appl. 49 (1983) 161–178.
[10] L. Verde-Star, Operator identities and the solution of linear matrix difference and differential equations, Stud. Appl. Math. 91 (1994) 153–177.
[11] L. Verde-Star, Taylor functionals and the solution of linear difference equations, in: Applications of Fibonacci Numbers, vol. 7, Graz, 1996, Kluwer, Dordrecht, 1998, pp. 449–462.
[12] L. Verde-Star, Solving linear differential-like equations, Amer. Math. Monthly 107 (2000) 213–226.
[13] L. Verde-Star, Functions of matrices, Linear Algebra Appl. 406 (2005) 285–300.